MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I.Memo No. 517 April, 1979

Some properties of discontinuities in the image irradiance equation

Anna R. Bruss

Abstract: The image irradiance equation is a first order partial differential equation. Part of this paper is a "comprehensive" guide to solving this kind of equation. The special structure of the image irradiance equation is explored in order to understand the relation of discontinuities in the surface properties and in the image intensities.

Acknowledgments: This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under the Office of Naval Research contract N00014-75-C-0643.

Table of contents

I. Motivation

The question whether reconstructing the shape of an opaque object by measuring the light it reflects is possible was raised in Horn's thesis [HO70]. We will refer to this as the *shape from shading* problem. He observed that under certain assumptions (which we will discuss in the next section), first order partial differential equations describe the relation between the brightness of a small patch of an object and the local surface normal. In other words we can determine the shape of an object by solving a first order partial differential equation (abbreviated in the following by (FO)PDE), also referred to as the *image irradiance equation.*

The literature about PDE's is extensive, but the emphasis is on higher order PDE's. As the majority of physical phenomena can be formulated as second order PDE's, these equations have been studied the most.

In this paper we are going to study FOPDE's. We try to summarize the various known results and describe methods for solving a given FOPDE. At the same time we keep in mind that the equations describe a physical situation and therefore their solutions have to make "sense". A major problem is that the mathematical literature deals nearly exclusively with equations and their solutions which are continuous and have continuous derivatives in all their variables. But "real" objects have edges and there the surface normals are discontinuous. Another problem are occluding contours in the pictures as there the partial derivatives of the function describing the surface are discontinuous.

It is intuitively "clear" that if the FOPDE is discontinuous in some of its variables, we can expect solutions which have discontinuous derivatives. The question arises if we can have such solutions also if the equation itself is continuous. Is it possible for example that an object has an edge without the

equation reflecting this fact. The answer is yes and we will examine why and how this can happen. We will show that in this case initial conditions are going to reflect the discontinuity.

In general, a PDE describes a class of processes and not a a particular instance of one of them. Consider as an example the Laplace equation

$$\Delta f = 0$$

where Δ denotes the Laplace operator and f a vectorfield. Then the PDE tells us that the field f has no sources and zero curl. But there are a lot of fields which fall in this category. Only when we specify some more conditions about f can we determine the unique solution of the equation.

We will study what kind of constraints are necessary to pin down a unique solution of a given image irradiance equation. An attempt will be made to find constraints which are accessible, i.e., which can be measured.

II. The *shape from shading* problem revisited

There are basically three components to this problem which we have to understand. They are the lightsource, the object and the viewer.

The exposure of film in a camera (for fixed shutter speed) is proportional to image irradiance, the flux per unit area falling on the image plane. Similarly, grey levels measured in a electronic imaging device are quantized measurements of image irradiance. It can be shown that image irradiance in turn is proportional to scene radiance, the flux emitted by the object per unit projected surface area per unit solid angle [HOS78]. The factor of proportionality depends on to details of the optical system including the effective f-number.

Scene radiance depends on the

1) surface material and its microstructure,
2) the distribution of light source, and,
3) the orientation of the surface.

Consider a viewer-oriented coordinate system with the viewer located far above the surface on the z-axis. If the objects imaged are small compared to their distance from the viewer, one can approximate the imaging situation by an orthographic projection,

$$x' = x(f/z_0) \qquad y' = y(f/z_0)$$

where (x',y') are the coordinates of the image of a point (x,y,z) made with a system of focal length f, when the viewer is at distance z_0 above the origin. We assume that $(x^2 + y^2 + z^2) \ll z_0^2$.

The orientation of a patch of the surface can be specified by given its gradient (p,q), where p and q are the first order partial derivatives of z with respect to x and y. For a particular surface material and a particular distribution of light sources, scene radiance will depend only on surface gradient. This function, $R(p,q)$, (or a contour representation in gradient space), is called the reflectance map.

If $L(x,y)$ in the scene radiance calculated form the observed image irradiance at the point (x',y') in the image then

$$R(p,q) = L(x,y) \tag{2.2}$$

where (p,q) is the gradient at the corresponding point on the object being imaged. This equation is called the *image irradiance equation*. It is clearly a first order partial differential equation since it involves only the partial derivatives p and q and the coordinates x and y.

A word of caution: we are not dealing with several issues like mutual

illumination, shadows and specularity.

III. Basics

For simplicity of exposition we will only deal with partial differential equations involving a function z of two variables x and y. It is more or less "obvious" how to generalize the results to functions of n variables. We will denote z_x and z_y - the partial derivatives of z with respect to x and y - by p and q respectively. Then the relation

$$F(x,y,z,p,q) = 0 \qquad (3.1)$$

is called a first order partial differential equation. A function z(x,y) is called a solution of (3.1) if in some region of the x-y plane the function and its derivatives satisfy the equation identically in x and y. Such a function is also called an integral surface.

When solving a partial differential equation we want to find the "general" solution which is a whole "set" of solutions. By imposing some additional constraints we can find the particular solution in which we are interested. Such constraints can be for example boundary conditions or initial values. In a later section we will state precisely what we mean by a "general" solution and what kind of constraints are necessary to pin down the desired solution.

The relation (3.1) is a linear PDE if it is linear in z, p and q with coefficients depending only on x and y and (3.1) is quasi-linear if it is linear in p and q with coefficients depending on x, y and z.

Unless otherwise stated we will assume that F, z and all relevant derivatives exist and are continuous.

IV. The quasi-linear first order PDE

We will first consider this special PDE as its geometric interpretation is rather clear and so the relevant method for solving it can be explained and understood easily. In this case the relation (3.1) can be rewritten as

$$a(x,y,z)p + b(x,y,z)q = c(x,y,z) \tag{4.1}$$

To rule out trivial cases we will further assume that

$$a^2 + b^2 \neq 0$$

We will try to find solutions to (4.1) given implicitly by

$$G(x,y,z) = 0 \tag{4.2}$$

Differentiating (4.2) with respect to x and y gives us

$$G_x + G_z z_x = 0 \quad \text{and} \quad G_y + G_z z_y = 0 \tag{4.3}$$

or equivalently

$$z_x = - G_x/G_z \quad \text{and} \quad z_y = - G_y/G_z \tag{4.3'}$$

Using (4.3') in (4.1) we get

$$a(x,y,z)G_x(x,y,z) + b(x,y,z)G_y(x,y,z) + c(x,y,z)G_z(x,y,z) = 0 \tag{4.4}$$

Note that in general (4.1) is a nonlinear PDE for the function $z(x,y)$ whereas (4.4) is a linear PDE for $G(x,y,z)$. We can interpret the coefficients $a(x,y,z)$, $b(x,y,z)$, $c(x,y,z)$ in (4.4) as the components of a vectorfield $\mathfrak{A} = \mathfrak{A}(x,y,z) = [a(x,y,z),b(x,y,z),c(x,y,z)]$.

Then we can rewrite (4.4) as

$$\langle\mathfrak{A}, \nabla G\rangle = 0 \tag{4.5}$$

where ∇G denotes the gradient of G and $\langle \, , \, \rangle$ the inner product of two vectors.

We know that at each point ∇G is perpendicular to the surface defined by $G(x,y,z)$ and the equation (4.5) tells us that $\mathfrak{A}$ is perpendicular to ∇G. Thus $\mathfrak{A}$ has to lie in the tangent plane of the integral surface defined by $G(x,y,z)$.

Let us introduce the notion of a fieldline of a vectorfield. By a fieldline

we understand a curve whose tangent at every point has the direction of the fieldvector there. Then an integral surface can be built up from fieldlines (called characteristics in this context) of the vectorfield $\mathfrak{A}$. To reiterate the previous statements: The tangent at each point of a characteristic has the same direction there as the vector $\mathfrak{A}$ and therefore by virtue of (4.5) the same direction as the tangent plane of the integral surface $G(x,y,z)$. This does not mean that each quasi-linear PDE has a single solution. Such a PDE only constrains the possible orientations of the tangent planes at each point to a one-parameter manifold. As (4.1) is linear in p and q all feasible tangent planes at every point of an integral surface intersect in a line which is called the Monge axis. Thus finding a solution to (4.1) means finding a surface which at each point has the direction of the Monge axis (the direction of the vector $\mathfrak{A}$) as its tangent direction.

Let us now describe a method of finding the characteristic curves, which can be written as functions of one parameter $x=x(s)$, $y=y(s)$ and $z=z(s)$. Then $\dot{\mathfrak{x}}(s)=[\dot{x}(s),\dot{y}(s),\dot{z}(s)]$ (where $\dot{}$ denotes differentiation with respect to s) has the same direction as $\mathfrak{A}$ and therefore the outer product of $\dot{\mathfrak{x}}(s)$ and $\mathfrak{A}$ has to be zero.

$$\begin{aligned} b\dot{z} - c\dot{y} &= 0 \\ c\dot{x} - a\dot{z} &= 0 \\ a\dot{y} - b\dot{x} &= 0 \end{aligned} \tag{4.6}$$

The relation (4.6) is normally written as

$$dx : dy : dz = a : b : c \tag{4.7}$$

The solutions of the equations (4.7) comprise a two-parameter family of curves in space (the characteristics). We know that only a one-parameter subset of them generate the solutions of the PDE . To find this subset we introduce an

arbitrary function between the two free parameters we get solving (4.7). So the general solution of (4.1) contains an arbitrary function of one parameter.

Now we are able to summarize all the previous stated results as follows: Each surface which is produced by a one-parameter family of characteristics is an integral surface. Conversely each integral surface is generated in such a fashion.

The first statement should be clear by now (otherwise this is not a "comprehensive" guide to first order PDE's). To understand the second, remark the following.

On each integral surface $z = z(x,y)$ the equations

$$\frac{dx}{ds} = a(x,y,z) \qquad \frac{dy}{ds} = b(x,y,z)$$

define a one-parameter family of curves: $x = x(s)$, $y = y(s)$, $z = z(x(s),y(s))$. Note that on such a curve

$$\frac{dz}{ds} = c(x,y,z)$$

as

$$\frac{dz}{ds} = \frac{\partial z}{\partial x}\frac{dx}{ds} + \frac{\partial z}{\partial y}\frac{dy}{ds} = az_x + bz_y = c$$

<u>example</u>:

$$F(x,y,z,p,q) = xp + yq - z = 0$$

Then the equations for the characteristics (4.7) are

$$dx : dy : dz = x : y : z$$

and have as their solution the two-parameter curves in space

$$y = C_1 x$$
$$z = C_2 x \qquad (*)$$

Now we introduce an arbitrary function w between the two constants: $w(C_1) = C_2$ to find a one-parameter subset of (*)

$$\frac{z}{x} = C_2 = w(C_1) = w\left(\frac{y}{x}\right)$$

solution of the PDE : $z = w(y/x)x$

or in parameter form : $y = C_1 x$ and $z = w(C_1)x$.

V. Method of characteristics for general first order PDE's

We want to apply similar methods as developed in section IV to the general first order PDE

$$F(x,y,z,p,q) = 0 \qquad (5.1)$$

To exclude trivial cases we will assume that

$$F_p^2 + F_q^2 \neq 0$$

Our goal is to transform the problem of finding a solution to (5.1) to the problem of integrating a set of ordinary differential equations. Again geometrical reasoning will help us find these equations.

Let us fix a point P with coordinates (x,y,z) on an integral surface. Then the quantities p and q are constrained by (5.1) to a one-parameter family of curves. [in other words: Once x, y and z are fixed, (5.1) is an equation for p and q. To write this equation in parameter form we only need one parameter.] As p and q determine the direction of the tangent plane at P, we have just established the fact that (5.1) constraints the feasible tangent planes to a one-parameter family. The envelope of the tangent planes is a conical surface and is

called Monge cone. This surface can have several sheets. Then "the considerations here refer merely to a suitable small range of tangent planes, e.g., a portion of a sheet of the cone where q can be expressed as a single-valued differentiable function of p [COHI62]." Each generator of this cone represents a possible direction of the tangent plane at P and is called a characteristic direction. Thus the integral surface has to "fit" into the field of Monge cones.

Recall now that in the quasi-linear case the Monge cone degenerated to the Monge axis, i.e. at every point the direction of the tangent plane was fixed. In that case we proceeded by finding characteristic curves which at every point had as their tangent direction the direction of the Monge axis there. We concluded that the integral surface is swept out by the characteristic curves. Actually we can do the same thing in the case of a general PDE, but we have to be a little more careful this time. First we find the curves which at every point have as their tangent direction the characteristic direction. Let such a curve (called focal curve) be given by $x(s)$, $y(s)$ and $z(s)$. Remember that we are looking for curves which sweep out the integral surface $z(x,y)$. In other words we want the functions $x(s)$, $y(s)$, $z(s)$, $p(s)$ and $q(s)$ to satisfy the PDE (5.1). The focal curves only determine $x(s)$, $y(s)$ and $z(s)$ and (5.1) gives us only a relation between p and q. So we are one equation short in order to determine p and q. We will obtain this equation by forcing the focal curves to lie on the integral surface. The focal curves, which also satisfy this last condition are called characteristic curves. Again the characteristic curves sweep out the integral surface.

Actually the problem we are concerned with is to find the integral surface. So we have to go the "opposite" way from what was described in the preceding paragraphs. We will first find a set of equations, called the characteristic

equations. A subset of the solutions of this set are the characteristic curves, from which the integral surface can be built up.

In the following paragraphs we are going to develop the technical "machinery" to find an integral surface of (5.1).

Our first task is to find the equation of the Monge cone. So let us fix a point (x,y,z). Then we can write p and q - which satisfy (5.1) - as functions of a parameter u and all feasible tangent planes at (x,y,z) can be expressed as

$$(Z-z) = (X-x)p(u) + (Y-y)q(u) \tag{5.2}$$

The envelope of the planes defined by (5.2) defines a conical surface with vertex at (x,y,z), the Monge cone. [aside: a conical surface is produced by moving a straight line which is fixed at one point along a curve]. We get the equation of the Monge cone by eliminating u from (5.2) and (5.3) which is obtained by differentiating (5.2) with respect to u.

$$0 = \frac{(X-x)\,dp}{du} + \frac{(Y-y)\,dq}{du} \tag{5.3}$$

Differentiating the PDE (5.1) with respect to the parameter u we obtain

$$0 = F_p\frac{dp}{du} + F_q\frac{dq}{du} \tag{5.4}$$

Assuming that neither dp and dq nor F_p and F_q vanish identically we get from (5.2), (5.3) and (5.4) that

$$\frac{X-x}{F_p} = \frac{Y-y}{F_q} = \frac{Z-z}{pF_p + qF_q} \tag{5.5}$$

By substituting all possible values for p and q (i.e. all values for p and q which satisfy (5.1)) we obtain all generators of the Monge cone at the point (x,y,z). The generators of the Monge cone at the different points of the integral surface define the tangent direction of the focal curves. Therefore the

focal curves have to satisfy the following differential equations.

$$\frac{dx}{ds} = F_p \qquad \frac{dy}{ds} = F_q \qquad \frac{dz}{ds} = pF_p + qF_q \tag{5.6}$$

Let $z = z(x,y)$ be an integral surface on which we also know p and q. Then the equations

$$\frac{dx}{ds} = F_p \qquad \frac{dy}{ds} = F_q \tag{5.7}$$

define a one-parameter family of curves.

On these curves

$$\frac{dz}{ds} = z_x\frac{dx}{ds} + z_y\frac{dy}{ds} \tag{5.8}$$

holds and using (5.7) in (5.8) we obtain

$$\frac{dz}{ds} = pF_p + qF_q$$

[aside: The above condition is called strip condition. "It states that the functions x(s), y(s), z(s), p(s), q(s) not only define a space curve, but simultaneously a plane tangent to it at every point. A configuration consisting of a curve and a family of tangent planes to this curve is called a strip [COHI62]."]

So (5.8) states that the curves defined by (5.7) are focal curves. Now we also require that a focal curve is embedded on an integral surface: ["By embedding we mean that in the neighborhood of the projection of a focal curve on the x-y plane z is a single-valued, twice continuous differentiable function of x and y [COHI62]."] If we differentiate the PDE (5.1) with respect to x and y we

get

$$F_p p_x + F_q q_x + F_z p + F_x = 0$$
$$F_p p_y + F_q q_y + F_z q + F_y = 0 \tag{5.9}$$

Using (5.7) and the fact that $p_y = q_x$ we get

$$\frac{dp}{ds} = p_x \frac{dx}{ds} + p_y \frac{dy}{ds} = p_x F_p + q_x F_q$$
$$\frac{dq}{ds} = q_x \frac{dx}{ds} + q_y \frac{dy}{ds} = p_y F_p + q_y F_q \tag{5.10}$$

Using (5.9) in (5.10) we get

$$\frac{dp}{ds} + F_z p + F_x = 0$$
$$\frac{dq}{ds} + F_z q + F_y = 0 \tag{5.11}$$

We can now summarize the previous results as follows: If a focal curve is embedded on an integral surface then along the curve the coordinates x,y,z and the quantities p and q satisfy the following five ordinary differential equations:

$$\frac{dx}{ds} = F_p \qquad \frac{dy}{ds} = F_q \qquad \frac{dz}{ds} = pF_p + qF_q$$
$$\frac{dp}{ds} = -(pF_z + F_x) \qquad \frac{dq}{ds} = -(qF_z + F_y) \tag{5.12}$$

Let us now consider the system (5.12) by itself, i.e. disregarding that we obtained it with a given integral surface in mind. Note first that $F(x,y,z,p,q)$ is constant along each solution of the system (5.12) as

$$\frac{dF}{ds} = F_p\frac{dp}{ds} + F_q\frac{dq}{ds} + F_z\frac{dz}{ds} + F_x\frac{dx}{ds} + F_y\frac{dy}{ds} =$$

$$= - F_p(pF_z + F_x) - F_q(qF_z + F_y) + F_z(pF_p + qF_q) + F_xF_p + F_yF_q = 0$$

Thus $F(x,y,z,p,q)$ = constant is a solution of (5.12). The system of equations (5.12) defines a four-parameter family of solutions. By imposing the additional constraint that the solutions of (5.12) also satisfy the PDE (5.1) we obtain a three-parameter subset of the solutions, the characteristic strip. "A space curve $x(s)$, $y(s)$, $z(s)$ bearing such a strip is called characteristic curve [COHI62]." We have already established the fact that a one-parameter subset of the three-parameter family of curves sweeps out the integral surface.

As characteristics depend on the solution, their range of influence cannot be determined in advance.

In the next section we will discuss the notion of a complete integral and then we will show how to choose the appropriate one-parameter subset.

So the problem of finding a solution to (5.1) is equivalent to integrating the system of five ordinary differential equations (5.12) which are also called the characteristic equations.

VI. General solution and complete integral

In the previous section we saw that each solution of a general first order PDE is swept out by a one-parameter family of curves. Thus the equation of an integral surface can be written as a function of the coordinates x and y and of an arbitrary function of one variable; such an equation is called the general solution.

Let us now assume for a moment that we have a solution $z = \phi(x,y,a,b)$ of the PDE which depends on two parameters. Then we say $\phi(x,y,a,b)$ is a complete integral if

$$D = \phi_{xa}\phi_{yb} - \phi_{xb}\phi_{ya}$$

is not equal to zero. This condition assures that ϕ really depends on two parameter, i.e. that there is no $\alpha = g(a,b)$ such that $\phi(x,y,a,b) = \phi(x,y,\alpha)$.

From the two-parameter family of planes defined by $\phi(x,y,a,b)$ we can choose a one-parameter subset by introducing an arbitrary function which relates a and b, e.g. set $b = w(a)$. Note that the family $\phi(x,y,a,w(a))$ is again a solution of the PDE . The following idea makes the concept of a complete integral significant: The envelope of the family $\phi(x,y,a,w(a))$ is again a solution of the PDE since at each point it touches a member of the family $\phi(x,y,a,w(a))$ i.e. a solution. Or conversely each point of the envelope is a solution of the PDE . We obtain the equation of the envelope by eliminating the parameter a from the two equations:

$$\begin{gathered} z = \phi(x,y,a,w(a)) \\ \phi_a(x,y,a,w(a)) + \phi_b(x,y,a,w(a))w'(a) = 0 \end{gathered} \tag{6.1}$$

Eliminating the parameter a from (6.1) yields an expression involving an arbitrary function w of one variable, which is a solution to the PDE and therefore we have found the general solution. We will now exhibit this fact analytically. By differentiating (6.1) with respect to x and y we get

$$\begin{aligned} z_x &= \phi_x + (\phi_a + \phi_b w'(a))a_x \\ z_y &= \phi_y + (\phi_a + \phi_b w'(a))a_y \end{aligned} \tag{6.2}$$

We know that $\phi(x,y,a,w(a))$ is a solution for any choice of the parameter a.

Using (6.1) (i.e. $\phi_a + \phi_b w'(a) = 0$) in (6.2) establishes the fact that for all x and y the values of z, z_x and z_y are the same as those of ϕ, ϕ_x and ϕ_y.

So if we know the complete integral of a given PDE we can obtain the general solution by just using the process of differentiation and elimination of parameters. (This later process can in practice be tedious or impossible, but is often not necessary, as by plugging in all different values for a, all solutions of the PDE are obtained.) The only problem with the above described method is that there is no easy way to find a complete integral. In the next section we will show that with the help of the characteristic equations we can find a complete integral.

The general solution does not comprise all solutions of a PDE . The envelope of the complete integral, the so called singular integral, is again a solution which cannot be obtained from the general solution. The equation of the singular integral, which does not contain any arbitrary elements, is found by eliminating the parameters a and b from the equations

$$\begin{aligned} z &= \phi(x,y,a,b) \\ \phi_a(x,y,a,b) &= 0 \\ \phi_b(x,y,a,b) &= 0 \end{aligned} \tag{6.3}$$

We have assumed all along that all eliminations are possible and that during the course of this process we obtain functions with continuous derivatives.

Actually we do not have to know the complete integral in order to find the singular solution. Note only that, for a complete integral, $F(x,y,\phi,\phi_x,\phi_y)$ vanishes identically for all choices of the parameters a and b. If we now differentiate the PDE with respect to a and b we get

$$F_\phi \phi_a + F_p \phi_{xa} + F_q \phi_{ya} = 0$$
$$F_\phi \phi_b + F_p \phi_{xb} + F_q \phi_{yb} = 0 \qquad (6.4)$$

As ϕ is a complete integral, $D = \phi_{xa}\phi_{yb} - \phi_{xb}\phi_{ya}$ is not equal to zero. Furthermore ϕ_a and ϕ_b are zero (equations (6.3)) and therefore we get the equation of the singular integral from (6.4) by eliminating p and q from

$$F_p = 0 \quad F_q = 0 \quad F = 0 \qquad (6.5)$$

[aside:we did not assume in this case that $F_p^2 + F_q^2 \neq 0$ as we did when obtaining the characteristic equation.]

If the PDE does not contain the function z(x,y) explicitly then there exist no singular solution as in this case the complete integral is of the form

$$z = \phi(x,y,a) + b$$

and the condition $\phi_b = 0$ can not be fulfilled.

VII. Method for finding the complete integral

In the previous sections we have shown two methods for finding the solutions to a given PDE . Now we will show how, with the help of the characteristic equations, we can find a complete integral and this will also then be a way to find a one-parameter subset from the four-parameter family of solutions of the characteristic equations.

First we have to discuss a special form of PDE called Pfaff's equation.

$$f(x,y,z)dx + g(x,y,z)dy + h(x,y,z)dz = 0 \qquad (7.1)$$

In the case when $h \equiv 0$ and f and g depend only on x and y (7.1) degenerates to an ordinary differential equation which is called an "exact" differential equation.

$$f(x,y)dx + g(x,y)dy = 0 \tag{7.2}$$

The equation (7.2) is called "total" if f and g satisfy the integrability condition

$$f_y(x,y) = g_x(x,y) \tag{7.3}$$

In the case of a total differential equation it is easy to find a solution to (7.2). On each simply connected region we can find a function H(x,y) such that

$$\frac{\partial H}{\partial x} = f(x,y) \quad \text{and} \quad \frac{\partial H}{\partial y} = g(x,y) \tag{7.4}$$

then

$$dH = f(x,y)dx + g(x,y)dy$$

and the equation $dH = 0$ is equivalent to (7.2).

Thus $H(x,y) =$ constant is a solution to (7.2) and the function H can be found by integrating (7.4).

In the case that (7.3) is not satisfied by the given equation (7.2) one can always introduce an "integration" factor $\mu(x,y)$ such that the equation

$$\mu f dx + \mu g dy = 0 \tag{7.5}$$

is total, i.e. that $(\mu f)_y = (\mu g)_x$. Or equivalently $\mu(x,y)$ has to be a solution of the following PDE which can be solved with the method of characteristics.

$$\mu(f_y - g_x) + \mu_y f - \mu_x g = 0$$

Now let us return to equation (7.1) which is again easy to solve if its left hand side is a total differential of a function H(x,y,z) i.e. if

$$f = \frac{\partial H}{\partial x} \qquad g = \frac{\partial H}{\partial y} \qquad h = \frac{\partial H}{\partial z} \tag{7.6}$$

Necessary for (7.6) to hold is that

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \qquad \frac{\partial g}{\partial z} = \frac{\partial h}{\partial y} \qquad \frac{\partial h}{\partial x} = \frac{\partial f}{\partial z} \tag{7.7}$$

In a simply connected area (7.7) is also sufficient for the existence of a

function $H(x,y,z)$ which satisfies (7.6) and can be calculated as

$$H(x,y,z) = \int_{(x_0,y_0,z_0)}^{(x,y,z)} (f dx + g dy + h dz) + C$$

where (x_0, y_0, z_0) is a fixed point. Clearly $H(x,y,z)$ = constant is a solution to (7.1).

In the case when (7.7) is not satisfied we again want to find an "integration" factor $\mu(x,y,z)$ such that the expression $\mu f dx + \mu g dy + \mu h dz$ is a total differential of a function. In comparison to Pfaff's equation in two variables it is not always possible to find such a factor.

Necessary for μ to exist is that the following equation holds:

$$f(g_z - h_y) + g(h_x - f_z) + h(f_y - g_x) = 0 \qquad (7.8)$$

It can also be shown (simple, but tedious) that in a simply connected region (7.8) is sufficient for (7.1) to possess a one-parameter family of solutions $H(x,y,z)$ = constant. We will show now how to construct such a function $H(x,y,z)$.

First let us consider the "abbreviated" equation

$$f(x,y,z)dx + g(x,y,z)dy = 0 \qquad (7.9)$$

This is a Pfaff's equation in the two variables x and y with z as a parameter. Thus we can always (eventually with the help of an integrating factor $\lambda(x,y,z)$) find a solution to (7.9):

$$\Phi(x,y) = u(x,y,z) = C \qquad (7.10)$$

Note that

$$\lambda f = \frac{\partial u}{\partial x} \quad \text{and} \quad \lambda g = \frac{\partial u}{\partial y} \qquad (7.11)$$

Now we define a function S depending on the three variables x,y and z by

$$S(x,y,z) = \lambda h - \frac{\partial u}{\partial z} \tag{7.12}$$

If we express - with the help of (7.10) - y as a function of x, u and z, we can redefine S as a function of x, u and z, i.e.

$$T(x,u,z) = S(x,y,z) \tag{7.13}$$

We will prove now that T is actually independent of x. Then we can find H(x,y,z) by solving another Pfaff's equation in the variables u and z. To prove that $\partial T/\partial x = 0$ we use equation (7.11) and (7.12) and obtain:

$$\begin{aligned}
\frac{\partial}{\partial x}(\lambda h - S) &= u_{xz} = \frac{\partial}{\partial z}(\lambda f) \\
\frac{\partial}{\partial y}(\lambda h - S) &= u_{yz} = \frac{\partial}{\partial z}(\lambda g) \\
\frac{\partial}{\partial x}(\lambda g) \quad &= u_{xy} = \frac{\partial}{\partial y}(\lambda f)
\end{aligned} \tag{7.14}$$

or (equation (7.14) written out in full)

$$S_x = h\lambda_x - f\lambda_z + \lambda(h_x - f_z) \tag{7.15}$$

$$-S_y = g\lambda_z - h\lambda_y + \lambda(g_z - h_y) \tag{7.16}$$

$$0 = f\lambda_y - g\lambda_x + \lambda(f_y - g_x) \tag{7.17}$$

Multiplying (7.15) by g, (7.16) by f and (7.17) by h and then adding up the three equations using condition (7.8) gives us

$$gS_x - fS_y = 0 \tag{7.18}$$

Differentiating (7.13) with respect to x and y gives us

$$S_x = T_x + T_u u_x$$
$$S_y = T_u u_y \tag{7.19}$$

If we combine now (7.11), (7.18) and (7.19) we obtain:

$$0 = gS_x - fS_y = gT_x + gT_u u_x - fT_u u_y =$$
$$= gT_x + \lambda fgT_u - \lambda fgT_u = gT_x$$

As $g \neq 0$ we can conclude that $T_x = 0$. Thus we can rewrite (7.13) as

$$S(x,y,z) = T(u,z)$$

The equation (7.1), after being multiplied by λ and with the expressions (7.11) and (7.12) used, reads now as follows:

$$\lambda(fdx + gdy + hdz) = u_x dx + u_y dy + (u_z + T)dz = 0$$

or equivalently

$$du + T(u,z)dz = 0 \tag{7.20}$$

This is again a Pfaff's equation in two variables which we know how to solve. Its solution is $\psi(u,z) = C$. Thus the solution to (7.1) is

$$H(x,y,z) = \psi(u(x,y,z),z) = C$$

which is one-parameter manifold.

Finally we are ready to describe a method for finding a complete integral of a general first order PDE $F(x,y,z,p,q) = 0$.

The basic idea is that we can interpret the total differential of a solution $z(x,y)$ of the PDE

$$dz = pdx + qdy \tag{7.21}$$

as Pfaff's equation in the variables x, y and z. To really do so we still have to express p and q as functions of x, y and z. Now assume that we can find two functions f and g such that if we set

$$p = f(x,y,z,a) \quad \text{and} \quad q = g(x,y,z,a) \tag{7.22}$$

(where a is an arbitrary constant) the PDE and condition (7.8) (with $h \equiv -1$)

$$fg_z - gf_z - f_y + g_x = 0 \tag{7.23}$$

is satisfied. Then the solution to (7.9) is a one-parameter manifold, but as we have already built in a parameter a into the equation (7.9) the solution contains two parameters and is the complete integral. So we have to solve the problem of how to find such functions f and g. We need two equations to do so. Let us assume that somehow a function $G(x,y,z,p,q)$ exists such that p and q (or equivalently f and g) can be expressed from the two equations

$$\begin{aligned} F(x,y,z,p,q) &= 0 \\ G(x,y,z,p,q) &= a \end{aligned} \tag{7.24}$$

(aside: thus $F_pG_q - F_qG_p \neq 0$).

Now we want to assure that f and g obtained in such a fashion satisfy (7.23) identically in the three variables x, y and z. Differentiating (7.24) with respect to x, y and z gives us:

$$\begin{aligned} F_x + p_xF_p + q_xF_q &= 0 \qquad & G_x + p_xG_p + q_xG_q &= 0 \\ F_y + p_yF_p + q_yF_q &= 0 \qquad & G_y + p_yG_p + q_yG_q &= 0 \\ F_z + p_zF_p + q_zF_q &= 0 \qquad & G_z + p_zG_p + q_zG_q &+ 0 \end{aligned} \tag{7.25}$$

After expressing p_z, q_z, p_y and q_x from (7.25) and plugging this expression into (7.23) we obtain the linear first order PDE for the function G

$$F_pG_x + F_qG_y + (pF_p + qF_q)G_z - (F_x + pF_z)G_p - (F_y + qF_z)G_q = 0 \tag{7.26}$$

We can solve (7.26) with the method of characteristics. The appropriate

system of characteristic equations is the same as the one for $F(x,y,z,p,q) = 0$

$$dx:dy:dz:dp:dq = F_p:F_q:(pF_p + qF_q):-(pF_z + F_x):-(qF_z + F_y) \qquad (7.27)$$

But we need now only one integral of (7.27) - which is independent of F and contains at least one of the variables p and q. Such an integral is our desired function G. There will always be such an integral as the solution of (7.27) comprise a four-parameter family:

$$v_i(x,y,z,p,q) = C_i \qquad i = 1,2,3,4$$

The v_i are independent and at least one of them must contain either p or q.

The just described method is due to Lagrange and Charpit. It has the advantage over the method of characteristics as described in section V in that we only need to find a single integral of (7.27) instead of finding the four-parameter family of curves.

VIII. Initial-value problem for linear and quasi-linear PDE's

Now that we know how to find the general solution of a given first order PDE we will attack the problem to determine the constraints with which a particular solution can be found.

We will consider the quasi-linear PDE

$$a(x,y,z)z_x + b(x,y,z)z_y = c(x,y,z) \qquad (8.1)$$

We now want to find the integral surface $z(x,y)$ which passes through a given curve C in space (in the literature referred to as Cauchy's problem). Clearly the following questions have to be answered:

1) What conditions on C are necessary such that this problem is solvable.

2) When is such a solution unique.

Let C be given by continuous differentiable functions of a parameter t: x(t), y(t), z(t). Furthermore we will assume that the projection of C on the x-y plane (later referred to as C_0) does not contain double points - (without this constraint we obtain surfaces with self intersections, i.e. z is not everywhere a single-valued function of x and y, which implies that along the line of intersection p and q are discontinuous) - and that $x_t^2 + y_t^2 \neq 0$. Now to construct a solution of the PDE which contains C we lay through each point of C a characteristic curve whose equations depend now on two parameters:

$$x = x(s,t) \quad y = y(s,t) \quad z = z(s,t) \tag{8.2}$$

Note that the functions x, y, z are still continuous differentiable. To get the equation of the integral surface we have to eliminate the parameters s and t from the equations (8.2), i.e. we have to express s and t in terms of x and y. A sufficient condition to do so is that the functional determinant Δ as specified in (8.3) does not vanish along the curve C.

$$\Delta = x_s y_t - y_s x_t \tag{8.3}$$

[aside: using the characteristic equations we can rewrite (8.3) as $\Delta = ay_t - bx_t$]

Thus if $\Delta \neq 0$ we can express z as a function of x and y and it is assured that C lies on the surface. The solution is also unique which follows from the following lemma:

> Each characteristic curve which has one point in common with an integral surface, lies completely on this surface.

The proof of this lemma is the uniqueness theorem for solutions of ordinary

differential equations.

The determinant Δ can be interpreted as the outer product of the two vectors

$$\xi_1 = \begin{pmatrix} x_s \\ y_s \end{pmatrix} \quad \text{and} \quad \xi_2 = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$$

which are the projections of the tangent and the characteristic direction on the x-y plane. In the special case when Δ vanishes along C these two directions coincide and we can deduce that C has one of the three properties listed below:

1) C is a characteristic curve.

2) C is the envelope of the characteristics. (called edge of regression)

3) C_0 is the envelope of the projections of the characteristics on the x-y plane.

Let us first discuss case 1). From $\Delta = 0$ we get

$$\frac{x_t}{a} = \frac{y_t}{b} \tag{8.4}$$

If we use x(t) and y(t) (from the equation for C) in z(x,y) then the following equation has to hold along C:

$$\frac{dz}{dt} = z_x\frac{dx}{dt} + z_y\frac{dy}{dt} = az_x + bz_y = c$$

Or equivalently we can say that C satisfies the characteristic equations and is therefore a characteristic curve. Obviously the solution of the PDE is not unique in this case. Actually in the case when C is a characteristic curve there exist infinitely many surfaces through C which satisfy the PDE . To see that just choose another curve C' along which Δ vanishes and which has a point P in common with C. Now to construct the solutions through C' one lays the characteristic through every point of C', in particular also through P, (i.e. the characteristic curve through P is C). Thus an integral surface through C'

contains C. In such a way infinitely many integral surface can be constructed which contain C. They meet along the characteristic curve and we can say that C is a branch curve.

One assumption we had made should be stressed here again: we were looking for solutions of the PDE which in the neighborhood of C are continuous and are continuous differentiable. It might be possible to find a solution z through C, along which Δ vanishes, without C being characteristic. These are the cases 2 or 3 as mentioned above. But then the derivatives of z are not continuous on C. Let us illustrate this problem with an example.

<u>example</u>

$$F(x,y,z,p,q) = 3(z - y)^2 p - q = 0$$

characteristic equations:

$$\frac{dx}{ds} = 3(z - y)^2 \qquad \frac{dy}{ds} = -1 \qquad \frac{dz}{ds} = 0 \qquad (*)$$

solution of (*) with initial values x_0, y_0, z_0:

$$x = (z_0 - y_0 + s)^3 + x_0 - (z_0 - y_0)^3$$
$$y = - s + y_0 \qquad (**)$$
$$z = z_0$$

We want the solutions (**) pass through C given by:

$$x = 0 \quad y = t \quad z = t$$

Note that C is not a characteristic.

So we set the initial values x_0, y_0, z_0 (i.e. x,y,z for s = 0) as

$$x_0 = 0 \quad y_0 = t \quad z_0 = t$$

With these initial values (**) becomes

$$x = s^3$$

$$y = -s + t$$

$$z = t$$

and the determinant Δ is in this case

$$\Delta = x_s y_t - x_t y_s = 3s^2$$

Thus along the curve C (i.e. s = 0) $\Delta = 0$

But there is a solution of the PDE which passes through C:

$$z = x^{1/3} + y$$

Note that $p = 1/3x^{-2/3}$ does not exist along C (as $x = 0$ there).

Now we will explain the previous example. To deduce from $\Delta = 0$ that C is a characteristic curve we used the relation:

$$\frac{dz}{dt} = z_x \frac{dx}{dt} + z_y \frac{dy}{dt} \tag{8.5}$$

But (8.5) is based on the following lemma in analysis:

Let G be a region in R^2, $f:G \to R$. If the partial derivatives at $c \in G$ exist and are continuous in a neighborhood of c then f is differentiable at c.

This lemma is not satisfied if C is an edge of regression or C_0 is the envelope of the projection of the characteristic curves on the x-y plane, (cases 2 and 3 in the above list).

In the case of a linear PDE we can make some more statements about the solution if Δ vanishes along C. We will show that in this case the integral

surfaces are cylindrical surfaces perpendicular to the x-y plane, i.e. that the solution is independent of z. The PDE is

$$a(x,y)p + b(x,y)q = c(x,y)$$

Recall that Δ is defined as

$$\Delta = x_s y_t - x_t y_s$$

Note that if Δ_s (i.e. the partial derivative of Δ with respect to s) and Δ vanish along C, then Δ vanishes everywhere. (proof: existence and uniqueness theorem for ordinary differential equations.)

Using the characteristic equations:

$$x_s = a \qquad (*)$$

$$y_s = b$$

we obtain for Δ_s

$$\Delta_s = a_s y_t + a y_{st} - b_s x_t - b x_{st} =$$

$$= a_s y_t + a b_t - b_s x_t - a_t b$$

Differentiating a and b with respect to s and t and using relations (*) we get

$$a_s = a_x a + a_y b$$

$$a_t = a_x x_t + a_y y_t$$

$$b_s = b_x a + b_y b$$

$$b_t = b_x x_t + b_y y_t$$

and then

$$\Delta_s = (a_x + b_y)\Delta$$

We want to express x as a function of y and z, i.e. $x = f(x,y)$ and so we assume that $y_s z_t - z_s y_t \neq 0$ along C. Then we will proceed showing that $f_z = 0$, which will prove that the integral surface z is a cylindric surface.

Differentiating x with respect to s and t we obtain

$$x_s = f_y y_s + f_z z_s$$

$$x_t = f_y y_t + f_z z_t$$

then

$$\Delta = f_z(z_s y_t - z_t y_s)$$

from which it follows

$$f_z = 0$$

IX. Initial-value problem for general first order PDE's

We have seen that the problem of finding a solution to a general first order PDE is equivalent to solving the system of characteristic equations. Again we are posing the question about what kind of constraints determine a solution uniquely. Clearly we need more than in the quasi-linear case, as now the the solutions to the characteristic equations form a three-parameter family of curves. So let C be a curve given by x(t), y(t), z(t) such that neither C nor its projection on the x-y plane have double points. Furthermore we have to specify p(t) and q(t) along C such that the condition

$$\frac{dz}{dt} = p\frac{dx}{dt} + q\frac{dy}{dt}$$

and the PDE (5.1) (i.e. F = 0) holds identically in t. We say that the function x(t), y(t), z(t), p(t), q(t) define an initial integral strip denoted by C_1. From now on the procedure is very similar to the one for solving the initial value problem for a quasi-linear PDE . So through every element of C_1 we lay a characteristic strip, which then can be written as x(s,t), y(s,t), z(s,t), p(s,t), q(s,t). To express the parameters s and t in terms of x and y we demand that

$$\Delta = x_s y_t - x_t y_s = F_p y_t - F_q x_t$$

does not vanish identically along the initial strip. Then z, p and q can be expressed in terms of x and y. We only have to make sure that p and q written in such a fashion are the partial derivatives of the integral surface z(x,y). Thus we have to show that the quantities U and V

$$\begin{aligned} U &= z_t - px_t - qy_t \\ V &= z_s - px_s - qy_s \end{aligned} \tag{9.1}$$

vanish identically. As we have assumed that $\Delta \neq 0$ we can deduce from (9.1) and

$$0 = z_t - z_x x_t - z_y y_t$$

$$0 = z_s - z_x x_s - z_y z_y$$

that $z_x = p$ and $z_y = q$. Recall now the characteristic equations:

$$\frac{dx}{ds} = F_p \qquad \frac{dy}{ds} = F_q \qquad \frac{dz}{ds} = pF_p + qF_q$$

Using the first two in the last we obtain

$$\frac{dz}{ds} = p\frac{dx}{ds} \quad q\frac{dy}{ds}$$

which implies $V \equiv 0$.

Now to prove $U \equiv 0$:

$$\frac{\partial U}{\partial s} = z_{st} - p_s x_t - px_{st} - q_s y_t - qy_{st} \tag{9.2}$$

$$\frac{\partial V}{\partial t} = z_{st} - p_t x_s - px_{st} - q_t y_s - qy_{st} \tag{9.3}$$

(9.2) - (9.3):

$$\frac{\partial U}{\partial s} - \frac{\partial V}{\partial t} = -(p_s x_t - p_t x_s + q_s y_t - q_t y_s) \tag{9.4}$$

Taking into account the characteristic equations and the fact that $V \equiv 0$ implies $\partial V/\partial t = 0$ we can rewrite (9.4) as

$$\frac{\partial U}{\partial s} = p_t F_p + q_t F_q + x_t F_x + y_t F_y + (p x_t + q y_t) F_z \tag{9.5}$$

But we know that the PDE $F = 0$ holds identically in s and t. Differentiating F with respect to t:

$$F_x x_t + F_y y_t + F_z z_t + F_p p_t + F_q q_t = 0 \tag{9.6}$$

(9.6) in (9.5):

$$\frac{\partial U}{\partial s} = - F_z U \, .$$

For any fixed t this is an ordinary differential equation for U as a function of s with the solution:

$$U(s) = U(0) e^{\int_0^s - F_z ds}$$

As by assumption $U(0)$ is zero, U vanishes everywhere.

To summarize the previous results: given a curve $x(t)$, $y(t)$, $z(t)$ along which we also know $p(t)$, $q(t)$ such that

$$\frac{dz}{dt} = p\frac{dx}{dt} + q\frac{dy}{dt}$$

$F(x(t),y(t),z(t),p(t),q(t)) = 0$ and $\Delta = F_p y_t - F_q x_t \neq 0$, then there exists an unique integral surface through the initial strip. We get a unique surface because the solution of the characteristic equations is uniquely determined by its initial values.

The exceptional case when $\Delta = 0$ along C_1 is analogous to the one discussed in the previous section: there are infinitely many integral surfaces through C_1 if

and only it is a characteristic strip. Again we can say that the characteristic curves are branch elements as on either side there can be another member of the family of solutions of the PDE while we are assured that along C the first derivatives are continuous.

If C_1 is only a focal strip along which $\Delta = 0$, then it might be still possible to find an integral surface z through it. But analog to the quasi-linear case, z will not have continuous derivatives.

The last case we will discuss is if C degenerates to a point P with coordinates (x_0, y_0, z_0). Then the strip condition is identically satisfied for all p_0 and q_0 which also satisfy the PDE, i.e., for all p_0 and q_0 which determine the feasible tangent planes in P. Thus we can write p_0 and q_0 as functions of a parameter t. If we plug the quantities x_0, y_0, z_0, $p_0(t)$, $q_0(t)$ in the PDE we obtain an integral surface which is in this case a conical surface with vertex at P. It is called the integral conoid of the partial differential equation at P.

It can be shown that the solution to the Cauchy problem can also be found by constructing the appropriate integral conoids. The integral surface is the envelope of the integral conoids whose vertices lie along a given curve C.

X. Summary

In the preceding chapters we dealt with FOPDE's which are continuous and have continuous derivatives in all their variables. We showed that the problem of finding a solution to a given FOPDE is equivalent to solving a system of ordinary differential equations. By imposing "suitable" initial conditions, a unique solution can be found.

But recall that our motivation for studying FOPDE's was to be able to solve

the image irradiance equation:

$$R(p,q) = L(x,y) \tag{10.1}$$

This equation has two properties which we are going to exploit. First the function z does not appear explicitly in the equation and therefore no singular integral surface exists. Second the equation can be written as the difference of two functions, where one depends only on p and q whereas the other depends solely on x and y. We will use this fact to deal with discontinuities.

Let us first review the case when (10.1) is a linear equation (note that (10.1) cannot be quasi-linear!), i.e;

$$p \pm q = L(x,y)$$

We assume that $L(x,y)$ is continuous and has continuous first derivatives. Let C be an initial curve given in parameter form: $x = x(t)$, $y = y(t)$ and $z = z(t)$. If the Δ (as defined in section VIII) does not vanish along C, then we have a unique solution. It is continuous and has continuous derivatives if C is a continuous differentiable function of the parameter t. "Any singularities of the initial data propagate in the x-y plane along the projection there of the relevant characteristic curve [GAR64]." This is not surprising as characteristic curves can be viewed as branch curves in which two integral surfaces meet. Note also that higher order derivatives may be discontinuous along characteristic curves.

In the case when C or its projection onto the x-y plane has double points, the integral surface has self intersections, i.e. z is not a single-valued function of x and y. This is not a case of interest for us, as we can "see" only one value of z. Note that a Mobius-strip is not a contradiction to the above assertions, as it is not an opaque object. In a case like that of a Mobius-strip mutual illumination and shadows are essential for deducing the shape of the object.

If Δ vanishes along C, then C is either a characteristic curve and we get

infinitely many solutions or we get a solution whose derivatives are not continuous along C. This second case is of special interest to us. Assume we could specify such a C. This would imply that there is for instance an edge which we do not see, as it is neither reflected in the equation itself nor in the initial data. But luckily this cannot happen. In the linear case the characteristic equations are:

$$\dot{x}_s = 1 \quad \text{and } y_s = \pm 1$$

Thus for Δ to vanish $x_t = y_t$ which is equivalent to saying that $x(t)$ and $y(t)$ have to satisfy the characteristic equations.

Now we will discuss the general image irradiance equation. Again we will assume that $R(p,q)$ and $L(x,y)$ are continuous and have continuous first derivatives. In deriving the system of characteristic equations we assumed that the integral surface z also has continuous second derivatives [we used this fact in equation [5.10] to deduce that $p_y = q_x$]. In [MY48, PLI54] it was shown that an integral surface can be build from characteristic strips also in the case when only the first derivatives are continuous.

We want the integral surface to pass through a curve C again given in parameter form by $x = x(t)$, $y = y(t)$ and $z = z(t)$. Then we can determine $p(t)$ and $q(t)$ along C by solving the two equations:

$$R(p(t),q(t)) = L(x(t),y(t)) \tag{10.2}$$

$$\frac{dz}{dt} = p(t)\frac{dx}{dt} + q(t)\frac{dy}{dt} \tag{10.3}$$

which we consider here to be algebraic (not differential) and solved over the reals. As (10.2) is nonlinear, we may get zero or several solutions for $p(t)$ and $q(t)$. Thus only for a unique determination of the roots of the equations (10.2) and (10.3) and the assumption that $\Delta \neq 0$ along C do we get a unique integral

surface. Recall that in the linear case two integral surfaces intersect smoothly (i.e. have the same tangent plane along their intersection) only along characteristic curves. In the general case this can also happen along the curve where the transition from one root of p and q to another takes place. To simplify the following discussion we will assume that p and q are given along C, such that equations (10.2) and (10.3) are satisfied. In other word the initial data is an initial strip denoted by C_1. Again discontinuities in the initial data propagate into the solution. It was shown [HA28], that if the initial data does not have continuous second derivatives, then the solution does not have continuous first derivatives. To understand this fact, recall the lemma of Schwarz:

Let G be a simply connected region in R^2, $f:G \to R$. If f_x, f_y and f_{xy} exist and are continuous, then f_{yx} exists and $f_{xy} = f_{yx}$.

Furthermore for p and q to be the partial derivatives of z, they have to satisfy the above lemma. Thus if the initial data does not have continuous second derivatives, then the lemma of Schwarz is not satisfied along C_1, which implies that p and/or q are discontinuous.

The case $\Delta = 0$ is analogous to the linear case. If C_1 is a characteristic strip, then we get infinitely many solutions. Again we pose the question whether it is possible to specify C_1 such that $\Delta = 0$ and C_1 is not characteristic. Then the derivatives of p and/or q would be discontinuous along C (which by the preceding remarks implies that the first derivatives are discontinuous). In that case C would be a focal curve, i.e. satisfying the equations

$$\frac{dx}{dt} = R_p \quad \frac{dy}{dt} = R_q \quad \frac{dz}{dt} = pR_p + qR_q \tag{10.4}$$

but

$$\frac{dp}{dt} \neq L_x(x(t),y(t)) \quad \text{and} \quad \frac{dq}{dt} \neq L_y(x(t),y(t)) \tag{10.5}$$

This cannot occur as we assumed that the function $L(x,y)$ is continuous and has continuous derivatives. Thus if $x(t)$ and $y(t)$ satisfy (10.4) they also have to satisfy (10.5) as L_x and L_y depend only on x and y.

The last case we have to be concerned is when Δ vanishes only along parts of C [LER57]. But then C cannot be a continuous differentiable function and the solution is going to have discontinuous derivatives.

We always assumed that C is not a closed curve, as this would either overdetermine the problem or make it inconsistent.

XI. Open questions

Very little is known about "singular" PDE's, i.e. equations which are not continuous in all their variables. Again some literature, e.g. [HAD28] can be found about singular second order PDE's.

In the case of the image irradiance equation we are concerned about singularities which occur in the equation itself. We can always assume that $R(p,q)$ is continuous in p and q. A case when $R(p,q)$ is not continuous is when the surface contains specularities, which can be represented by delta-functions. Then a completely different approach has to be taken to solve the equation. But in a lot of cases $L(x,y)$ is going to be discontinuous. As an example we use the following situation: looking through an electron microscope on a half-sphere lying on a flat surface which can be formulated as:

$$p^2 + q^2 = \frac{x^2 + y^2}{1 - (x^2 + y^2)}$$

Along the contour $x^2 + y^2 = 1$ the equation is discontinuous. In this case the contour is called an occluding contour. Informally we can say that if $L(x,y)$ is discontinuous then the derivatives of the integral surface are going to be discontinuous. By specifying the solution along the curve C' where the discontinuity occurs, we can still find an integral surface. As z does not appear explicitly in the equation, only the projection of C' onto the x-y plane can be found directly from the equation. Rigorous proofs of these claims are going to appear in a subsequent paper.

We always assumed that the integral surface is continuous. Discontinuities in z correspond to gaps and for these kind of situations the image irradiance equation is inadequate.

Another open problem is how to deal with discontinuities in the image intensity gradient, i.e. discontinuity in L_x and/or L_y. To be able to integrate the characteristic equations we had to assume that the first derivatives of L and R are continuous. We conjecture that the case of discontinuities in the first derivatives can be treated in an analogous fashion as the cases of discontinuities in the functions themselves.

The image irradiance equation describes some of the important features of the "real" world. There is evidence that human beings deduce a lot of information about the shape of an object by looking at its contours and registering the "grey" levels. We discussed how these two pieces of evidence are tied together in a single partial differential equation..

Acknowledgments

I want to thank Professor Berthold Horn for writing section II and for his helpful comments concerning this paper. Furthermore I am grateful to Eric Grimson for proofreading an earlier draft and to Ken Forbus for listening.

Bibliography

[COHI62] Courant, R. and D. Hilbert, *Methods of Mathematical Physics,* Vol.II. Interscience Publishers, 1962, pp.1-147.

[COP75] Copson, E.T., *Partial Differential Equations,* Cambridge University Press, 1975.

[DES72] Descovich, J., "Höhere Mathematik", *Lecture Notes,* Technische Universität Wien, 1972.

[DIR72] Dirschmid, H.-J., "Mathematik A3", *Lecture Notes,* Technische Universität Wien, 1972.

[EP62] Epstein, B., *Partial Differential Equations: An Introduction,* McGraw Hill Book Company Inc., 1962.

[GAR64] Garabedian, P.R., *Partial Differential Equations,* John Wiley & Sons, Inc., 1964.

[GOGU73] Golubitsky, M. and V. Guillemin, *Stable Mappings and Their Singularities,* Springer Verlag, 1973.

[HA28] Haar, A., "Über die Eindeutigkeit und Analyzität der Lösungen Partieller Differentialgleichungen," *Atti del congr. intern. dei mat.,* Bologna, Vol.3, 1928, pp.5-10.

[HAD64] Hadamard, J., *La Théorie des Equations aux Derivées Partielles,* Editions Scientifiques, Pekin, 1964.

[HO70] Horn, B.K.P., "Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from one View," *MIT, Project MAC TR-79,* 1970.

[HO75] Horn, B.K.P., "Obtaining Shape from Shading Information," *The Psychology of Computer Vision,* P.H. Winston (ed.), McGraw Hill Book Company Inc., 1975, pp.115-155.

[HO75] Horn, B.K.P., "The Facts of Light," *Artificial Intelligence Laboratory, Working Paper 97,* 1975.

[HOS78] Horn, B.K.P. and R.W. Sjoberg, "Calculating the Reflectance Map," *Artificial Intelligence Laboratory, A.I.Memo 498,* 1978.

[JOH78] John, F., *Partial Differential Equations,* Springer Verlag, Third Edition, 1978.

[LER57] Leray, J., "Uniformisation de la Solution du Problème Linéaire Analytique de Cauchy près de la Variéte qui Porte les Données de Cauchy," *Bull. Soc. Math. France,* 85, 1957, pp.389-429.

[MY48] Myškis, A., "The uniqueness of the Solution of Cauchy's Problem, *Uspehi Matem. Nauk (N.S.)*, 3, No.2(24), 1948, pp.3-46. (in russian); an english abstract can be found in *Mathematical Reviews*, 1949, pp.302-303.

[PLI54] Pliś, A., "Characteristics of Nonlinear Partial Differential Equations," *Bull. Acad. Polon. Sci.* Cl.III., 2, 1954.

[SMI68] Smirnow, W.I., *Lehrgang der Höheren Mathematik*, Vol.4, VEB Deutscher Verlag der Wissenschaften, Berlin 1968, pp.259-325.

[WU66] Wunderlich, W., *Darstellende Geometrie I.*, Bibliographisches Institut, Mannheim, Wien, Zürich, 1966.